Detection of GPT-4 Generated Text in Higher Education: Combining Academic Judgement and Software to Identify Generative AI Tool Misuse

Year: 2023 | Citations: 110 | Authors: Mike Perkins, Jasper Roe, Darius Postma, James McGaughran, Don Hickerson British Ur

---

Abstract

This study explores the capability of academic staff assisted by the Turnitin Artificial Intelligence (AI) detection tool to identify the use of AI-generated content in university assessments. 22 different experimental submissions were produced using Open AI's ChatGPT tool, with prompting techniques used to reduce the likelihood of AI detectors identifying AI-generated content. These submissions were marked by 15 academic staff members alongside genuine student submissions. Although the AI detection tool identified 91% of the experimental submissions as containing AI-generated content, only 54.8% of the content was identified as AI-generated, underscoring the challenges of detecting AI content when advanced prompting techniques are used. When academic staff members marked the experimental submissions, only 54.5% were reported to the academic misconduct process, emphasising the need for greater awareness of how the results of AI detectors may be interpreted. Similar performance in grades was obtained between student submissions and AI-generated content (AI mean grade: 52.3, Student mean grade: 54.4), showing the capabilities of AI tools in producing human-like responses in real-life assessment situations. Recommendations include adjusting the overall strategies for assessing university students in light of the availability of new Generative AI tools. This may include reducing the overall reliance on assessments where AI tools may be used to mimic human writing, or by using AI-inclusive assessments. Comprehensive training must be provided for both academic staff and students so that academic integrity may be preserved.